SUPPORTING STATEMENT

Part B

Use of Deliberative Methods to Enhance Public Engagement in the Agency for Healthcare Research and Quality's (AHRQ's) Effective Healthcare (EHC) Program and Comparative Effectiveness Research (CER) Enterprise

June 15, 2012

Agency for Healthcare Research and Quality (AHRQ)

Table of Contents

B. Collections of Information Employing Statistical Methods

1. Respondent universe and sampling methods

The goal of our sampling strategy is to obtain a sufficiently large sample and allocate sample members to different deliberative methods. In doing so, we seek adequate statistical power to test each of our research questions while maximizing external and internal validity to the extent possible. The five conditions discussed in this statement (four deliberative methods plus the control condition) are named as:

> Method A: Brief Citizens' Deliberation (BCD),
> Method B: Online Deliberative Polling® (ODP),
> Method C: Community Deliberation (CD),
> Method D: Citizens' Panel (CP), and
> Condition X: Control.

Respondent Universe

The goal of public deliberation is to obtain informed opinions from the *lay public*. It is based on the principle that persons affected by a given decision should have the opportunity to contribute their views regarding the decision (Webler, 1995). In public deliberation this principle is operationalized by selecting samples of participants who are "broadly representative" of the affected public (Rower & Frewer, 2000).

The study will take place in four purposively selected geographic regions. This is because three of the deliberative methods require in-person participation in groups at a common location. We selected the four locations to achieve racial/ethnic, socioeconomic, and geographic diversity; these are: Chicago, IL; Washington, DC; Raleigh-Chapel Hill-Durham, NC; and Sacramento, CA. We selected two very large, highly urbanized areas with relatively large African American and Hispanic populations to ensure we will have adequate samples of ethnic minority populations (Chicago and Washington DC) and two locations that are moderate-sized cities surrounded by rural areas within easy driving distance to increase access to non-urban residents. We will draw samples roughly representative of the local populations in each of the four areas. An important strength of this approach is that it minimizes participant burden, as participant travel is local. This has the additional likely benefit of increasing the range of individuals willing to participate in the study, and thus the representativeness of participants. Inclusion criteria for the study population are: English-speaking, ages 18 and over, and residency in one of the four selected locations.

Internet access. The sample will be further limited to individuals with access to the Internet. This criterion is imposed because two of the four deliberative methods require Internet access, and random assignment of the study sample to deliberative methods requires comparability of the sample on key factors across the four deliberative methods and the control group. This approach increases our confidence that any observed differences among the deliberative methods and controls are due to the variations in

method, and not to characteristics that may be systematically associated with having Internet access, enhancing internal validity by keeping participants' average Internet experience equal across the methods.

We have chosen the significant improvement in internal validity derived from limiting our sample to those with Internet access at the expense of a loss in external validity resulting from this strategy. Our judgment is based in part on data that show that persons without any internet access represent a small and rapidly decreasing segment of the U.S. population. The proportion of households connecting to the Internet at home using broadband grew from 9.2% in 2001, to 63.5% in 2009 (U.S. Department of Commerce, 2010). The current goal is for 98% of Americans to have access to high-speed wireless Internet services within five years (Office of the Secretary, 2011)

Although our sampling strategy will exclude persons with no internet access at all, it will include persons along the entire spectrum of internet usage from infrequent, low-skilled users to frequent, skillful users. We will measure internet sophistication with our pre-intervention survey and control for it in the statistical models.

Sampling Strategy

Our sampling approach will enable us to obtain samples that: (1) represent the local population's demographic distribution, and (2) provide adequate numbers of participants from AHRQ-defined priority populations, including participants who are Hispanic and bilingual in English and Spanish, and (3) have access to the Internet.

Participants will be recruited from the databases maintained by professional recruiting firms operating in the four study locales. AIR, AHRQ's contractor for this study, will contract with a local recruitment firm in each of the four geographic locations. The following section briefly describes the recruitment firms and recruiting approach.

The Roles of Recruitment Firms

North Carolina area, First in Focus. In the North Carolina area, AIR will retain First in Focus, a firm specializing in qualitative studies that is located in the Research Triangle region of North Carolina. This firm will recruit participants, provide the specified quota of priority populations with special emphasis on non-urban participants, facilitate recruits' participation, and process incentives.

Washington D.C. area, Shugoll Research Inc. The study will utilize Shugoll Research, Inc., a recruitment and marketing research agency in the Washington, D.C. metropolitan area. This firm will oversee recruitment of urban participants, facilitate recruits' participation, and administer incentives.

Chicago, Illinois area, FocusPointe Global. FocusPointe Global is a national recruitment and marketing research firm with a field office in Chicago. This firm focuses on urban populations and maintains a large database of participants from which to sample. FocusPointe Global will be recruiting with a specific emphasis on

Hispanic bilingual and African American women participant. This firm will also facilitate recruits' participation and administer incentives.

Sacramento, California area, Opinions Sacramento. Opinions Research focuses its work in the Sacramento area and maintains a large database of participants in the area. This firm will conduct recruitment with a specific emphasis on Hispanic bilingual and non-urban participants. This firm will also facilitate recruits' participation, and administer incentives.

We will use a quota sampling method based on the local demographic information provided by the U.S. Census Bureau. The sampling will occur in three steps in each of the four locations:

Step 1: We will define the geographic areas (i.e., counties) where study participants will be recruited from for each of the four locations.

Step 2: We will calculate the number of individuals required for each stratum or category to ensure that the distributions of gender, age, race, and Hispanic ethnicity in our sample reflect the local demographic distribution.

Step 3: We will randomly select members from each stratum until we obtain the required sample size within each stratum.

Based on current data from the U.S. Census Bureau,[1] the target sample size for each stratum is listed in Exhibit 1.

Exhibit 1. Target general population sample

	Target sample size based on combined populations in four areas
Total Population	1,296
Female	667
Aged 65 years and over	132
Race	
Black or African American	358
White and Other Race	938
Hispanic/Latino (any race)	183
Non-Urban	67

---

[1]   U.S. Census Bureau, Current Population Survey (CPS) and CPS, School Enrollment and Internet Use Supplement, October 2009.

Randomization

Within each of the four geographic locations, selected participants will be randomly assigned to the five experimental conditions (four deliberative methods or control). In a multi-treatment randomized study, it is typical to recruit a sample of participants who are willing and able to participate in any of the treatment options and then randomize the subjects across those options and the control condition. This approach is consistent with maximizing internal validity, a priority for this study. However, a major limitation of the typical randomization approach for this study is the disparate burden placed on participants by the deliberative methods, because the demands of the methods vary considerably. This difference introduces the potential for selection bias, because participants who are able and willing to participate in lower burden or online methods might differ significantly in their personal characteristics (e.g., age) or in some unobserved ways from those who are able and willing to participate in higher burden or in-person methods (See Supporting Statement A, Section A1 for a description of the methods). Because of the variation in burden and deliberation mode, the pool of individuals who will find all four methods acceptable (and thus be available for randomization using the typical approach) will most likely be small and not be representative of the populations we hope to represent. Thus, the typical randomization approach would likely be very expensive and/or have poor external validity.

One alternative is to draw separate random samples with their own control groups for each method. However, because of the variation in demands of the methods, the resulting samples would likely differ substantially from each other, threatening internal validity. Because of the inherent selection bias, this approach would limit our ability to compare specific deliberative methods to others (see Exhibit 5).

To address this genuine dilemma, we will use a two-step randomization procedure designed to maximize the pool of potential willing participants, minimize the cost of recruitment and yield internally and externally valid comparisons among specific deliberative methods (see Exhibit 5). The concept is that we will make these relevant comparisons among the methods (as listed in Exhibit 5) using only the subset of sample members who were willing to accept assignment to the methods being compared to each other. This approach minimizes selection bias in each comparison. The procedures for accomplishing this are explained below.

Step 1

Potential participants from the recruitment firms will be provided descriptions of the five experimental conditions and asked to indicate, on the electronic form or on the phone, which of the four deliberative methods they are willing to participate in. They will be asked to select as many of the methods as they are interested in and would take part in; they will be required to select at least two of the four methods to be eligible for the study. The instructions will further explain that they will be randomly assigned to one of their selected methods or a control condition, where they will be asked to read materials but not convene or take part in discussions. Excluding subjects who would only select one

4

method is necessary because we would not be able to use those persons when comparing one method to another.

As illustrated in Exhibit 2, this procedure will establish up to 11 pools of participants defined by all possible 3- to 5- way combinations of the four methods, where A through D represent the four methods and X represents the control group.

Exhibit 2. Potential Participant Pools Following Participant Selection of Two, Three, or Four Methods

Participants select 2 methods	Participants select 3 methods	Participants select 4 methods
6 combinations: ABX ACX ADX BCX BDX CDX	4 combinations: ABCX ABDX ACDX BCDX	1 combination: ABCDX

Step 2

In the second step, we will randomly assign each participant to one of his or her selected methods or the control condition. In Exhibit 3, we illustrate how randomization might work for our first four recruits. Recruit 1 will accept A and C, and Recruit 2 will accept B and D. Recruit 1 will be randomized to A, C, or X. Recruit 2 will be randomized to B, D, or X. Recruit 3 would be randomized to A, C, D, or X. Recruit 4 would be randomized to C, D, or X.

Recruits 3 and 4 would be eligible for the sample that is used to compare C and D, because both of them said they would accept assignment to C and D, even though they did not all choose the same combination of methods. However, they will be included in the C: D contrast only if they are actually randomized to C or D. This approach minimizes selection bias when making the pair-wise contrasts, because the analysis is limited to the pool of participants who find both methods acceptable.

Exhibit 3. Example Randomization Matrix of First Four Recruits as Demonstration

Example Recruit	Method A: Brief Citizens' Deliberation (BCD)	Method B: Online Deliberative Polling® (ODP)	Method C: Community Deliberation(CD)	Method D: Citizens' Panel (CP)
1	X		X	
2		X		X
3	X		X	X
4			X	X

Extending the illustration beyond the four recruits in Exhibit 3, eligibility for the contrast between methods A and B will be limited to participants who chose any combination that includes A and B—AB, ABC, ABD, ABCD—and the contrast will be made between the subset of those persons who were randomly assigned to either A or B. This approach ensures that each estimate includes all sample members for whom the respective treatment contrast is relevant.

Sample Allocation

Exhibit 4 shows the target sample allocation for the entire study. The minimum detectable effect size (MDES) for each of our research hypotheses, based on these target sample sizes across the five experimental conditions, is shown in the following section on Analysis and Statistical Power Considerations.

Exhibit 4. Sample Allocation Targets across the Five Experimental Conditions

Deliberative Method		Number of Participants and Groups
Brief Citizens' Deliberation: A	# participants per group	12
	# groups	24
	Total sample	288
Online Deliberative Polling®: B	# participants per group	12
	# groups	24
	Total sample	288
Community Deliberation: C	# participants per group	12
	# groups	24
	Total sample	288
Citizens' Panel: D	# participants per group	24
	# groups	4
	Total sample	96
Deliberative Groups Subtotal	# groups	76
	Total	960
Control: X	Total sample	336
Total Deliberation plus Control	**# groups**	**76**
	Total sample	**1,296**

We expect up to 30% attrition from the time of recruitment to the time the deliberative sessions are held; that is, we expect approximately 70% of respondents who say they will come to the deliberative sessions to participate. This estimate is based on team members' personal experience with deliberative groups, but it is also consistent with other group research for which participants are recruited (e.g., focus groups). In order to achieve the desired total sample size indicated in Exhibit 4, we will oversample by 30%, recruiting 1,680 participants to achieve the desired sample of 1,296 participants.

Analysis and Statistical Power Considerations

Answering Research Question 1 (see Supporting Statement, Section A16) calls for pair-wise comparisons of each method with the control group. We will answer Research Question 1 in two steps. First, we will compare each deliberative method to the control group to establish its individual effectiveness compared to education alone. Then we will pool deliberation participants from all effective methods to compare deliberation to control and establish the overall effect of deliberation. To make the individual comparison between a deliberative method and the control group, we will use all the participants who are assigned to either the pertinent method or the control group. For the pooled contrast between all the deliberative methods and control, we will use all the participants who are randomly assigned to one of the five experimental conditions.

Research Question 2 (see Supporting Statement, Section A16) addresses how burden changes the effectiveness of methods (i.e., knowledge of and attitudes toward the deliberative topics, and selected measures listed under quality of deliberative discourse and implementation in Section A16). We will conduct pooled comparisons of the two lower burden methods (BCD + ODP) vs. the two higher burden methods (CD + CP) to evaluate the impact of burden. We will also compare each of the lower burden methods with the pooled higher burden methods: BCD vs. (CD + CP) and ODP vs. (CD + CP). Finally, we will compare the methods with the greatest difference in burden, the lowest (BCD) vs. the highest (CP).

Research Question 3 (see Supporting Statement, Section A16) asks whether specific methods are more effective than other ones. We will conduct a comparison of methods that differ by mode: BCD (exclusively in-person) vs. ODP (exclusively on-line); both of which are low burden methods. We will also compare CD vs. CP to ascertain whether the most intensive method (CP) is substantially different in effectiveness from the less burdensome – but still relatively intensive – CD method.

Minimum Detectable Effect Size for Research Hypotheses

Based on information obtained from previous studies, we have chosen sample size targets to detect *moderate effect* sizes for the most important research hypotheses. With the current design, we estimate that the minimum detectable effect size (MDES) for the most important comparisons ranges from 0.21-0.40.

Exhibit 5 lists the minimum detectable effect size (MDES) with 80% power for each of our research hypotheses. The power calculations, conducted using Optimal Design Software (Spybrook, Raudenbush, Congdon, & Martinez, 2009), are based on the following assumptions:
1. The outcome measure is a continuous variable;
2. There are 12 persons in each deliberative group;
3. The intra-class correlation is 0.03 within each deliberative group[2];

_____

[2] The study participants will deliberate in groups (i.e., clusters), which is likely to make participants' deliberation experience and attitudes within the same group more similar to each other than they would be if deliberation were an individual endeavor and the participants were independent of each other. We have

4. The Type I error rate is 5% for two-tailed tests

Although the effect sizes we can detect in tests involving CP are larger than with our other methods, these are the tests where we expect to find the greatest differences in pair-wise comparisons. So, although the effect size for CP vs. control is 0.40 as compared to BCD, ODP, or CD vs. control at 0.27, there should be a greater difference between CP and control, because it is by far the most intensive deliberation experience.

Further, although it is (almost) always preferable to increase power, we have selected our strategy so that we can detect a medium effect size – one likely to be visible to the naked eye of a careful observer, as per Cohen (1992) – as what we need for evaluating choices between methods. By this criterion, if we do not detect a difference between CP and CD (MDES=0.40-0.73), we would be supported in pooling these methods. We expect that ultimately, small effect size will be trivial relative to the cost differences between the methods. Therefore, our focus is on detecting the moderate effect sizes that will be meaningful for future decisions about how to use public deliberation.

Exhibit 5. Statistical Power Estimates for Various Comparisons

Research Question	Minimum Detectable Effect Size (MDES)	
	Lower Bound*	**Upper Bound***
1. Is public deliberation more or less effective than education only.		
BCD vs. Control	0.27	
ODP vs. Control	0.27	
CD vs. Control	0.27	
CP vs. Control	0.40	
All methods combined vs. Control**	0.21	
2. How does burden change effectiveness?		
(BCD+ODP) vs. (CD+CP)	0.22	
BCD vs. CP	0.40	0.54
BCD vs. (CD+CP)	0.26	
ODP vs. (CD+CP)	0.26	
3. Are specific deliberative methods more effective than other ones?		
BCD vs. ODP	0.28	0.33
CD vs. CP	0.40	0.73

*Lower and upper bounds reflect maximum and minimum assumptions, respectively, about how many persons randomized to the method will be eligible for the comparison, based on our limited pilot experience. (Eligibility is determined by participants' willingness to participate in both methods involved in a comparison, reflecting participant choice rather than investigator assignment.)

---

taken this clustering (i.e., the intra-class correlation coefficient or ICC) into account when estimating the minimum detectable effect sizes in Exhibit 5.

As shown in Exhibit 5, to answer Research Question 1 ("***Is public deliberation more or less effective than education only?***"), we will be able to detect an effect as small as 0.27 when comparing BCD, ODP, or CD to the control group; and an effect of 0.40 when comparing CP to the control group. In addition, when we pool deliberation participants from all effective methods to compare deliberation to control, we will be able to detect an effect as small as 0.21, depending on the number of methods included in the pooled comparison.

To answer Research Question 2 ("***How does burden change effectiveness?***"), we will conduct pooled comparisons of the lower and higher burden methods. We can detect an effect as small as 0.22 in this comparison, specifically between BCD + ODP (the pooled lower burden methods) vs. CP + CD (the pooled higher burden methods. In comparing the least intensive method (BCD) with the most intensive (CP), we can detect an effect of 0.40-0.54.

To answer Research Question 3 ("***Are specific deliberative methods more effective than other ones?***"), we can detect an effect of 0.28-0.33 when comparing BCD (exclusively in-person method) with ODP (the exclusively online method) to examine the impact of mode. Further, we can detect an effect of 0.40-0.73 when comparing the two higher burden methods (CD vs. CP).

Evidence for Anticipated Effect Size
Our sample allocation targets are predicated on the assumption of finding a moderate effect for differences in pre-post change scores between experimental and control groups. The literature on effect sizes that can be expected from this design in studies of public deliberation is very limited. Two studies of deliberation (Farrar et al., 2010; Kim et al., 2010) have measured change in attitudes using multiple outcome measures and found pre-post effect sizes in the 0.17 to 0.69 range. A third study using a randomized control trial (RCT, Min 2007) compared a post-intervention measure of knowledge between experimental and control arms and found an effect size of 1.67 when comparing in-person deliberation to control and 2.05 when comparing Internet deliberation to control.

Illustration of Effect
Like these studies, our surveys contain knowledge and attitude questions. For knowledge questions, where each question is answered either correctly or incorrectly, we use the total number of correct answers among all knowledge questions as the overall knowledge score. An example from our Attachments C and E: Knowledge and Attitudes Pre- and Post-test Survey is:

For a new medicine to be approved for use in the United States, medical research results have to show that:
   a) The new medicine works better than medicines already approved.
   b) The new medicine is effective.
   c) The new medicine is approved in other countries.
   d) Don't know

Assuming the magnitude of our intervention effect is similar to Min (2007), we expect persons on average in the control group will answer 5 out of 8 questions correctly, while on average persons in the deliberative groups will answer 7.1 correctly. The score difference of 2.1 points is equivalent to the standardized effect size of 1.67 obtained by Min, a large effect based on Cohen's criteria[3].

Most of our attitude questions are scored on a 5-point scale. For example:

How important is it that people ask their doctors about medical research results related to their health problem?
1. Not important at all
2. Not important
3. No opinion
4. Important
5. Very Important

If we assume that the magnitude of the intervention effect (i.e., pre-post change of response to this question) and the variance of responses across study participants are similar to those obtained by Farrar (2010), we expect a mean score change of .71 points on a 5-point scale. In other words, we expect that an average score on our attitude question of 3 ('no opinion') before deliberation will shift to a score of 3.71 (closer to 'important') after the deliberation, a meaningful shift. From the statistical power perspective, the pre-post change of 0.71 scale points is equivalent to the standardized effect size of 0.69 that Farrar et al. reported, a moderate to large effect based on Cohen's criteria.

2.   Procedures for the Collection of Information

Knowledge and Attitudes surveys (Attachments C and E)

Objectives. The primary purpose of these surveys is to measure the effect of the deliberation on knowledge of and attitudes toward the deliberative topics. We will compare pre- and post-deliberation knowledge and attitude measures for each of the contrasts discussed above.

Participants. The surveys will be conducted with a 30% oversample of both deliberative methods participants and control subjects (n=1,680). The oversample is necessary to guard against the expected attrition of 30% between the time respondents consent and agree to participate in the study and the time of the actual implementation of the deliberative sessions. This estimate is based on our experience in recruiting for and implementing similar events.

Procedures. As described, recruitment for the study will be conducted by a local recruitment firm from each of the four geographic locations. Recruitment firms will call

---

[3] According to Cohen (1992), an effect size of 0.2 to 0.3 is regarded as small effect, an effect size around 0.5 is regarded as moderate effect and an effect size equal to or larger than 0.8 is regarded as large effect.

the selected sample of participants from their database and ask participants screening questions. Participants who agree to take part in the study will go on to receive, in electronic format, the full informed consent text (Attachment A). Participants will be asked to (1) select the deliberative methods they agree to take part in (per the selection procedures described above), (2) consent to completing the surveys that are part of the evaluation of the methods, and (3) consent to having their participation in the deliberative methods recorded by audio-recordings using procedures that will not identify individual participants. When the online consent process is complete, participants will be asked to complete the online baseline Knowledge and Attitudes survey. Participants will be given the option of completing the survey at the time of consent or return later to complete the survey. We will send email reminders to participants who do not complete the survey at the time of consent. The follow-up Knowledge and Attitudes survey will be administered online within a week of participation in the deliberative methods. The online survey will be administered to all sampled participants regardless of whether or not they actually participate in the deliberative sessions, which will enable us to conduct a non-response analysis. The follow-up survey will be administered to controls in a staggered period to correspond to the deliberative methods implementation schedule to minimize temporal bias.

Deliberative Experience Survey (Attachment F)

<u>Objectives.</u> As mentioned, the four deliberative methods vary in duration, mode, educational materials, and other characteristics described in Supporting Statement A. We expect that these differences will affect participants' experiences of and satisfaction with the deliberative process. To test this expectation, we will administer a post-intervention survey with persons who take part in the deliberative methods to obtain their reports about and ratings of the experience to determine participant preferences for each method. In the event that differences among methods in other outcomes, such as change in knowledge, attitudes, and beliefs, are small, participant preferences for characteristics of the method could be determining factors in deciding among methods to implement.

<u>Participants.</u> The survey will be administered to all individuals taking part in the deliberative methods (n=960). Control group members will not be surveyed.

<u>Procedures.</u> For the four in-person deliberative methods, a pen and paper survey will be administered prior to leaving the facility at which the deliberative method is held. For the online method, it will be administered online at the immediate conclusion of the final deliberative session.

Audio- and video-recording transcript review

<u>Objective.</u> To collect qualitative data to summarize and report the results of the deliberations on the issue, the *use of research evidence in healthcare decision-making.*

<u>Participants.</u> All deliberation participants. Control group members will be excluded.

Procedure. Audio- and video-recordings of all deliberative sessions will be transcribed and cleaned (e.g., correction of typographical errors, mis-identified speakers, inaccurate acronyms). Qualitative data will include transcripts and notes from team debriefings that will be conducted immediately following all deliberative sessions.

We will define and apply a coding system for the transcripts of the deliberative sessions. This coding system will be applied using qualitative analysis software to facilitate the retrieval, reduction, and analysis of data. The data will be coded by assigning labels to the units, clustering the codes into categories and hypothesizing about relationships among the categories. This process facilitates the identification of patterns and themes that could explain differences in outcomes.

3. Methods to Maximize Response Rates and Deal with Nonresponse

Participant recruitment. Local recruitment firms will be used to recruit participants and collect informed consent. AIR will administer the online Knowledge and Attitude pre-test and post-test surveys. Participants recruited by recruitment firms have higher response rates than population-based sample selection methods such as random digit dialing because the respondents have previously expressed interest in taking part in research projects. This reduces the number of persons that need to be contacted in order to achieve the desired sample size, which reduces the total burden. The recruitment firms collect demographic data on their members that can be used to achieve the desired participant mix (e.g., ensure participation by under-represented groups), which also reduces the number of people we need to contact and enables us to conduct non-response analysis.

Knowledge and Attitudes surveys (Attachments C and E). AIR will administer the online Knowledge and Attitude pre-test and post-test surveys. To maximize the response rate on these surveys, they will be administered online to make it convenient for participants to access and complete. Participants will be provided a payment of $25 for completion of each survey (pre-test and post-test). Compensation for time, effort, and inconvenience is typical for randomized controlled trials such as the one we are conducting and will improve the response rate for both the pre- and post-intervention surveys. We intend to administer both the pre- and post-intervention surveys to all 1,680 sampled persons, regardless of whether they participate in the deliberative sessions. This strategy will improve the quality of the nonresponse analysis by providing additional data on which to compare respondents and nonrespondents. It will also enable us to conduct an "intent-to-treat" analysis (i.e., including both respondents and nonrespondents in the analytic statistical models), which compensates for the potential biasing effects of attrition. Our analysis will compare the findings from the intent-to-treat sample with the findings from the sample limited to those who participated to help us understand the impact of attrition on our findings.

Deliberative Experience Survey (Attachment F). To maximize the response rate on this survey, it will be administered at the immediate conclusion of the final deliberative session of each method. Participants will not receive a payment specifically for this survey; however, they will receive a combined payment for their participation in the

deliberative session and completion of the immediate post-survey. The survey will be administered only to those sample members who complete the deliberation process to which they were assigned (n=960). It will not be administered to control group members.

Audio- and video- recordings transcript review. The transcript review will capture data from 100% of the 960 deliberation participants.

4. Test of Procedures or Methods to be Undertaken

Pilot of deliberative methods and qualitative analysis

We conducted a pilot test of the deliberative methods, implementing all four deliberative methods with fewer than 10 participants each. The goals of the pilot were to ensure that the various components of the study are implemented effectively. These components include: recruitment and randomization procedures, educational materials, use of experts, technology tools, facilitation, and administration of the Deliberative Experiences Survey.

All surveys were cognitively tested with fewer than 10 participants each in the development phase. Cognitive testing consists of 1-on-1 interviews, which identify cognitive challenges that may present barriers to the effective use of the instruments. We also conducted usability testing of the online surveys with fewer than 10 participants to be sure they are programmed correctly.

5. Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

Name (Affiliation)	Telephone Number	Email
Johannes Bos (AIR)	650-843-8110	jbos@air.org
Kristin Carman (AIR)	202-403-5090	kcarman@air.org
Kirsten Firminger (AIR)	919-918-4507	kfirminger@air.org
Steven Garfinkel (AIR)	919-918-2306	sgarfinkel@air.org
Dierdre Gilmore (AIR)	650-843-8139	dgilmore@air.org
Jessica Waddell Heeringa (AIR)	202-403-5947	jheeringa@air.org
Susan Heil (AIR)	301-592-2227	sheil@air.org
Maureen Maurer (AIR)	919-918-2308	mmaurer@air.org
Marilyn Moon (AIR)	301-592-2101	mmoon@air.org
HarmoniJoie Noel	202-403-5779	hnoel@air.org
Grace Wang (AIR)	650-843-8191	gwang@air.org
Amy Windham (AIR)	301-592-2165	awindham@air.org
Manshu Yang (AIR)	919-918-2312	myang@air.org
Stirling Bryan (consultant)	604-875-4776 (Canada)	Stirling.Bryan@ubc.ca
Todd Davies (Symbolic Systems Program, Stanford)	650-723-4091	davies@stanford.edu
James Fishkin (Center for Deliberative Democracy, Stanford)	650-723-4611	Fishkin@stanford.edu
Marge Ginsburg (Center for Healthcare Decisions)	916-851-2828	Ginsburg@chcd.org
Marthe Gold (consultant)	212-650-7794	goldmr@med.cuny.edu
Ela Pathak-Sen (consultant)	0-145-222-6206 (UK)	ela@commotionuk.com
Alice Siu (Center for Deliberative Democracy, Stanford)	650-724-1301	asiu@stanford.edu
Shoshanna Sofaer (consultant)	646-660-6815	Shoshanna.Sofaer@baruch.cuny.edu

References

1. Cohen, J. (1992). A power primer. *Psychological Bulletin,* 112 (1), 155–159.

2. Farrar, C., Fishkin, J. S., Green, D. P., List, C., Luskin, R. C., & Paluck, E. L. (2010). Disaggregating deliberation's effects: An experiment within a deliberative poll. *British Journal of Political Science,* 40(2), 333–347.

3. Kim, S. Y., Uhlmann, R. A., Appelbaum, P. S., Knopman, D. S., Kim, H. M., Damschroder, L. et al. (2010). Deliberative assessment of surrogate consent in dementia research. *Alzheimer's Dement,* 6(4), 342–350.

4. Min, S.-J. (2007). Online vs. face-to-face deliberation: Effects on civic engagement. *Journal of Computer-Mediated Communication,* 12(4), 1369–1387.
5. Office of the Press Secretary. (2011, February 10). *President Obama details plan to win the future through expanded wireless access.* Retrieved from http://www.whitehouse.gov/the-press-office/2011/02/10/president-obama-details-plan-win-future-through-expanded-wireless-access.

6. Rowe, G., & Frewer, L. J. (2000). Public Participation Methods: A Framework for Evaluation. *Science, Technology & Human Values, 25*(1), 3-29.

7. Spybrook, J., Raudenbush, S. W., Congdon, R., & Martínez, A. (2009). Optimal Design for Longitudinal and Multilevel Research: Documentation for the "Optimal Design" Software. University of Michigan, Ann Arbor, MI.

8. U.S. Department of Commerce, Economics and Statistics Administration and National Telecommunications and Information Administration. (2010). *Exploring the digital nation: home broadband internet adoption in the United States* Retrieved from http://www.esa.doc.gov/Reports/exploring-digital-nation-home-broadband-internet-adoption-united-states.

9. Webler, T. (1995). "Right" discourse in citizen participation: an evaluative yardstick. In O. Renn, T. Webler & P. Wiedemann (Eds.), *Fairness and Competence in Citizen Participation: Evaluating Models for Environmental Discourse* (pp. 35-86). Dordrecht: Kluwer Academic Publishers.